

ECS Configuration Change Request

Page 1 of 1 Page(s)

1. Originator Todd Edmands	2. Log Date: 1/28/03	3. CCR #: 03-0054	4. Rev: -	5. Tel: 303-735-2418	6. Rm #: 376B	7. Dept. NSIDC
8. CCR Title: System Monitoring						
9. Originator Signature/Date Todd Edmands /s/ 1/27/03			10. Class II	11. Type: CCR	12. Need Date: 2003-2-28	
13. Office Manager Signature/Date R. Erickson /s/ 1/27/03			14. Category of Change: Initial ECS Baseline Doc.		15. Priority: (If "Emergency" fill in Block 27). Routine	
16. Documentation/Drawings Impacted (Review and submit checklist): 6A 922-TDN-044			17. Schedule Impact: None		18. CI(s) Affected:	
19. Release Affected by this Change: 6A		20. Date due to Customer:		21. Estimated Cost: None - Under 100K		
22. Source Reference: <input type="checkbox"/> NCR (attach) <input type="checkbox"/> Action Item <input checked="" type="checkbox"/> Tech Ref. <input type="checkbox"/> GSFC <input type="checkbox"/> Other:						
23. Problem: (use additional Sheets if necessary) The system n0sas01 has had numerous CPU panics caused by hardware and software faults and occasionally the system is unable to return to a normal operational state without having to run manual fsck checks on the system drive. This problem has caused significant down time when failures happen and the system administration staff is not available (weekends). When the system is unavailable, ingested data cannot be written to the datapool.						
24. Proposed Solution: (use additional sheets if necessary) Install Sun Disk Suite Product which is included with Solaris that will allow the root drive of n0sas01 to be mirrored, increasing the availability of the system. Having a mirrored root drive will allow automatic consistency checks between the mirrored root file systems and inconsistencies will be resolved allowing the system to return to a normal run level in case of a system panic. The /data partition should also be mirrored or at a minimum changed to a journaled file system. Mirroring the filesystem will also provide faster recovery in case a root drive fails.						
25. Alternate Solution: (use additional sheets if necessary) Change local non qfs or swap partitions to journaled filesystems by adding the logging mount option in the /etc/vfstab file. This solution will prevent the problems of having fsck issues prevent the system from returning to an operational state, but does not create a highly available environment since disk mirror maintains a second copy of the file system.						
26. Consequences if Change(s) are not approved: (use additional sheets if necessary) Decreased system availability, and loss of data available to the data pool service. This server has had numerous CPU panics since being added to our inventory and occasionally inconsistencies on the root file system prevent the system returning to a normal run level. When this system is unavailable, ingested data is not written to the data pool filesystem made available by n0sas01..						
27. Justification for Emergency (If Block 15 is "Emergency"):						
28. Site(s) Affected: <input type="checkbox"/> EDF <input type="checkbox"/> PVC <input type="checkbox"/> VATC <input type="checkbox"/> EDC <input type="checkbox"/> GSFC <input type="checkbox"/> LaRC <input checked="" type="checkbox"/> NSIDC <input type="checkbox"/> SMC <input type="checkbox"/> AK <input type="checkbox"/> JPL <input type="checkbox"/> EOC <input type="checkbox"/> IDG Test Cell <input type="checkbox"/> Other						
29. Board Comments: DAAC has not implemented this solution. New CCR 03-0191 supercedes this CCR from NSIDC.			30. Work Assigned To:		31. CCR Closed Date:	
32. EDF/SCDV CCB Chair (Sign/Date):		Disposition: Approved App/Com. Disapproved Withdraw Fwd/ESDIS ERB Fwd/ECS				
33. M&O CCB Chair (Sign/Date): G. G. Gavigan, Jr. /s/ 02/04/03		Disposition: Approved App/Com. Disapproved Withdraw Fwd/ESDIS ERB Fwd/ECS /s/ 3/18/03 GGG.				
34. ECS CCB Chair (Sign/Date):		Disposition: Approved App/Com. Disapproved Withdraw Fwd/ESDIS ERB Fwd/ESDIS				

ADDITIONAL SHEET

CCR #: **Rev:** **Originator:** Todd Edmands

Telephone: 303-735-2418 **Office:** 376b

Title of Change: Mirror Root Disk

Requirements

- Two hard drives to mirror
- Add DiskSuite from the Solaris 8 installation media
- Restructure existing Swap partition, free 10MB of space for a slice 3 metadb file
- Create Metadb partitions on both drives (10MB)
- Restructure existing partition on second drive to match the first drive.
- configure metadevices
- reboot

----- CURRENT SYSTEM CONFIGURATION -----
System:N0sas01

Current Drive Usage:

swapfile	dev	swaplo	blocks	free
/dev/dsk/c0t0d0s1	32,1	16	4099424	4099424

FORMAT

AVAILABLE DISK SELECTIONS:

0. c0t0d0 <SUN18G cyl 7506 alt 2 hd 19 sec 248>
/pci@1f,4000/scsi@3/sd@0,0
1. c0t1d0 <SUN18G cyl 7506 alt 2 hd 19 sec 248>
/pci@1f,4000/scsi@3/sd@1,0

disk 0

Total disk cylinders available: 7506 + 2 (reserved cylinders)

Part	Tag	Flag	Cylinders	Size	Blocks
0	root	wm	0 - 1738	3.91GB	(1739/0/0) 8194168
1	swap	wu	1739 - 2608	1.95GB	(870/0/0) 4099440
2	backup	wm	0 - 7505	16.86GB	(7506/0/0) 35368272
3	unassigned	wm	0	0	(0/0/0) 0
4	unassigned	wm	0	0	(0/0/0) 0
5	unassigned	wm	0	0	(0/0/0) 0
6	unassigned	wm	2609 - 5216	5.86GB	(2608/0/0) 12288896
7	home	wm	5217 - 7505	5.14GB	(2289/0/0) 10785768

disk 1

Total disk cylinders available: 7506 + 2 (reserved cylinders)

Part	Tag	Flag	Cylinders	Size	Blocks
0	root	wm	0 - 55	128.84MB	(56/0/0) 263872
1	swap	wu	56 - 111	128.84MB	(56/0/0) 263872
2	backup	wu	0 - 7505	16.86GB	(7506/0/0) 35368272
3	unassigned	wm	0	0	(0/0/0) 0
4	unassigned	wm	0	0	(0/0/0) 0
5	unassigned	wm	0	0	(0/0/0) 0
6	usr	wm	112 - 7505	16.61GB	(7394/0/0) 34840528
7	unassigned	wm	0	0	(0/0/0) 0

Current File Systems on n0sas01

df -k

Filesystem	kbytes	used	avail	capacity	Mounted on
/dev/dsk/c0t0d0s0	4032142	1362698	2629123	35%	/
/proc	0	0	0	0%	/proc
fd	0	0	0	0%	/dev/fd

```

mnttab          0    0    0  0% /etc/mnttab
swap           2800584    16 2800568    1% /var/run
swap           524288    16 524272    1% /tmp
/dev/dsk/c0t0d0s6      6050182    119 5989562    1% /data
qfs1             2657279232 286401600 2370877632    11% /datapool
/dev/dsk/c0t0d0s7      5309706     9 5256600    1% /export/home

```

Current Swap partition on n0sas01

```

# swap -l
swapfile      dev swaplo blocks free
/dev/dsk/c0t0d0s1 32,1    16 4099424 4099424

```

Current /etc/vfstab configuration

```

#device      device      mount      FS      fsck  mount mount
#to mount    to fsck      point      type    pass  at boot options
#
#/dev/dsk/c1d0s2 /dev/rdisk/c1d0s2 /usr      ufs     1     yes  -
fd - /dev/fd fd - no -
/proc - /proc proc - no -
/dev/dsk/c0t0d0s1 - - swap - no -
/dev/dsk/c0t0d0s0 /dev/rdisk/c0t0d0s0 / ufs 1 no -
/dev/dsk/c0t0d0s6 /dev/rdisk/c0t0d0s6 /data ufs 2 yes -
/dev/dsk/c0t0d0s7 /dev/rdisk/c0t0d0s7 /export/home ufs 2 yes -
swap - /tmp tmpfs - yes size=512m
n0ins01:/var/mail - /var/spool/mail nfs - yes bg
qfs1 - /datapool samfs - yes stripe=0,writebehind=1152

```

----- REPARTITION First DRIVE -----

Overview: 10MB of disk space is needed for the metadb file on the first drive. Use the last 10MB of swap space to create an area for the state database.

Boot to single user mode and change the swap partition to the second drive.

```

/dev/dsk/c0t0d0s1 - - swap - no -
to
/dev/dsk/c0t1d0s6 - - swap - no -
REBOOT

```

reallocate the swap partition on the first drive to create 10mb of space for s3.

```

current value
1 swap wu 1739 - 2608 1.95GB (870/0/0) 4099440
3 unassigned wm 0 0 (0/0/0) 0

```

new values

```

Part Tag Flag Cylinders Size Blocks
1 swap wu 1739 - 2603 1.94GB (865/0/0) 4099440
3 unassigned wm 2604- 2608 10MB (5/0/0) 0

```

Verify that slice 3 has 10MB of space.

Reset the swap partition back to /dev/dsk/c0t0d0s1

REBOOT

----- REPARTITION SECOND DRIVE -----

Use the format utility to change the drive partitioning of c0t1d0s* to match c0t0d0s* for all partitions that will be mirrored. The second drive should now match the first drive's partition layout.

```

Part Tag Flag Cylinders Size Blocks
0 root wm 0 - 1738 3.91GB (1739/0/0) 8194168
1 swap wu 1739 - 2603 1.95GB (865/0/0) 4099440
2 backup wm 0 - 7505 16.86GB (7506/0/0) 35368272
3 unassigned wm 2604- 2608 10MB (5/0/0) 0
4 unassigned wm 0 0 (0/0/0) 0
5 unassigned wm 0 0 (0/0/0) 0
6 unassigned wm 2609 - 5216 5.86GB (2608/0/0) 12288896
7 home wm 5217 - 7505 5.14GB (2289/0/0) 10785768

```

----- Disk Mirroring Background -----

Root drives are never mirrored or put under Veritas version 2.x control because it is possible to encapsulate the root drives and to then mirror the drives. However in event of a system crash, the filesystem may become corrupt and

having a mirror will potentially cause the mirror to become corrupt too.

----- INSTALL SOFTWARE -----

Sun Disk Suite is included with the solaris operating system and does not require a license to be purchased.

Install Sun Solstice DiskSuite packages SUNWmd, SUNWmdr, SUNWmdu and SUNWmdg.
You do not need to install the Log Daemon package(SUNWmdn).

The DiskSuite packages have moved around a bit with each release of Solaris. With Solaris 8, DiskSuite moved to the "Solaris 8 Software" cdrom number two, in the EA directory. Starting the Solaris 9, DiskSuite is now included with the operating system. At the time of this writing, Solaris 8 is the most commonly deployed version of Solaris, so we'll use that as the basis for this example. The steps are basically identical for the other releases.

1. After having completed the installation of the Solaris 8 operating system, insert the Solaris 8 software cdrom number two into the cdrom drive. If volume management is enabled, it will automatically mount to /cdrom/sol_8_401_sparc_2 (depending on the precise release iteration of Solaris 8, the exact path may differ in your case):
2. Change to the directory containing the DiskSuite packages:
cd /cdrom/sol_8_401_sparc_2/Solaris_8/EA/products/DiskSuite_4.2.1/sparc/Packages
3. Add the required packages (we're taking everything except the Japanese-specific package):
pkgadd -d .

The following packages are available:

- 4 SUNWmdg Solstice DiskSuite Tool
(sparc) 4.2.1,REV=1999.11.04.18.29
- 2 SUNWmdja Solstice DiskSuite Japanese localization
(sparc) 4.2.1,REV=1999.12.09.15.37
- 3 SUNWmdnr Solstice DiskSuite Log Daemon Configuration Files
(sparc) 4.2.1,REV=1999.11.04.18.29
- 4 SUNWmdnu Solstice DiskSuite Log Daemon
(sparc) 4.2.1,REV=1999.11.04.18.29
- 5 SUNWmdr Solstice DiskSuite Drivers
(sparc) 4.2.1,REV=1999.12.03.10.00
- 6 SUNWmdu Solstice DiskSuite Commands
(sparc) 4.2.1,REV=1999.11.04.18.29
- 7 SUNWmdx Solstice DiskSuite Drivers(64-bit)
(sparc) 4.2.1,REV=1999.11.04.18.29

Select package(s) you wish to process (or 'all' to process all packages). (default: all) [?,??,q]: 4 6 7 8 9 10

*** IMPORTANT NOTICE ***

This machine must now be rebooted in order to ensure sane operation. Execute
shutdown -y -i6 -g0
and wait for the "Console Login:" prompt.

eject cdrom # shutdown -y -i6 -g0

4. Once the system reboots, apply any DiskSuite patches. At the time of this writing, the latest recommended DiskSuite patch available from sunsolve.sun.com is 108693-14 (DiskSuite 4.2.1). Note that the patch installation instructions require that a reboot be performed after the patch is installed
5. Proceed to the process of mirroring the operating system.

----- CREATE STATE DATABASES -----

How to Create Initial State Database Replicas From Scratch (Command Line)

When you run the metadb(1M) command for the first time, the system displays a warning message that no state database replicas exist for this host. Ignore this message. It appears only when you create the state database replicas for the first time.

The -a and -f options of metadb are used together to create the initial state database replicas. Five initial state database replicas, one on each of five slices, are created. By spreading the state database replicas across controllers, you can increase metadevice performance and reliability. The metadb command checks that the replicas are active, as indicated by the -a flag.

----- Mirror the operating system -----

Configuring DiskSuite as follows:

- * Create multiple copies of the State Database on slice 3 of each of the mirrored disks. In this case, we create four state database on each slice 6.

```
cd /usr/sbin
./metadb -af -c 4 /dev/dsk/c0t0d0s3
./metadb -af -c 4 /dev/dsk/c0t1d0s3
```

./metadb -i

(this command will display your state database info)

* Add the following to the /etc/vlm/md.tab file.

Be sure to make the changes appropriate for your environment.

----- Mirror the operating system -----

Mirror the root devices.

The high-level steps in this procedure are:

- Using metainit(1M) -f to put the root (/) slice in a single slice (one-way) concat (submirror1)
- Creating a second concat (submirror2)
- Using metainit(1M) to create a one-way mirror with submirror1
- Running the metaroot(1M) command
- Running the lockfs(1M) command
- Rebooting
- Using metattach(1M) to attach submirror2
- Recording the alternate boot path

Root Mirror MetaDevices

cd /usr/sbin

metainit -f d11 1 1 c0t0d0s0

d11: Concat/Stripe is setup

metainit d12 1 1 c0t1d0s0

d12: Concat/Stripe is setup

metainit d10 -m d11

d10: Mirror is setup

metaroot d10

lockfs -fa

reboot

...

metattach d10 d12

d10: Submirror d12 is attached

ls -l /dev/rdisk/c0t0d0s0

lrwxrwxrwx 1 root root 88 Feb 8 15:51 /dev/rdisk/c1t3d0s0 ->

../devices/pci@1f,4000/scsi@3/sd@0,0:a,raw

ls -l /dev/rdisk/c0t1d0s0

lrwxrwxrwx 1 root root 55 Jul 5 2002 /dev/rdisk/c0t1d0s0 ->

../devices/pci@1f,4000/scsi@3/sd@1,0:a,raw

* Wait for reync to finish before proceeding.

You can check the status via: /usr/opt/SUNWMd/sbin/metastat | more

----- NVRAM CONFIGURATION -----

Optional:

Add a device alias to the Open Boot Prompt NVRAM.

The system might already have the appropriate disk aliases already set up (i.e., "disk" and "disk1"). However, it is beneficial to add the following as devalias because they are more intuitive and this leaves the original devalias commands intact, should you ever need to go back to them:

{0} ok nvedit

0: devalias rootmirror0 /pci@1f,4000/scsi@3/sd@0,0:a

1: devalias rootmirror1 /pci@1f,4000/scsi@3/sd@1,0:a

{0} ok nvstore

{0} ok setenv use-nvramrc? true

{0} ok setenv boot-device rootmirror0 rootmirror1

{0} ok setenv diag-device rootmirror0 rootmirror1

-----MIRROR OTHER PARTITIONS IFNEEEDED-----

The high-level steps in this procedure are:

- Identifying the slice that contains the existing file system to be mirrored
- Using metainit(1M) -f to put the mounted file system's slice in a concat/stripe (submirror1)
- Creating a second concat/stripe (submirror2)
- Using metainit(1M) -m to create a one-way mirror with submirror1
- Unmounting the file system
- Editing the /etc/vfstab file so that the file system refers to the mirror
- Remounting the file system
- Using metattach to attach submirror2

```
# metainit -f d21 1 1 c0t0d0s6
d21: Concat/Stripe is setup
# metainit d22 1 1 c0t0d0s6
d22: Concat/Stripe is setup
# metainit d20 -m d21
d20: Mirror is setup
```

(Edit the /etc/vfstab file so that /usr references the mirror)

```
# reboot
```

```
...
```

```
# metattach d20 d22
```

```
d20: Submirror d22 is attached
```

```
# reboot if necessary
```